

# Domain-specific Design of Patient Classification in Cancer-related Cachexia Research

**Tiziana Margaria**

University of Limerick and Lero  
Limerick, Ireland

Alexander Wickert

University of Potsdam  
Potsdam, Germany

Anna-Lena Lamprecht

Utrecht University  
Utrecht, The Netherlands



UNIVERSITY  
of  
LIMERICK  
OILScoil Luimnigh



TRINITY  
COLLEGE  
DUBLIN



O'É Gaillimh  
NUI Galway



UCC  
University College Cork, Ireland  
Coláiste na hOllscoile Corcaigh



Maynooth  
University  
National University  
of Ireland Maynooth

# Motivation

Cachexia is a **complex wasting syndrome** associated with a marked detrimental effect upon life quality and survival in patients with **cancer, chronic obstructive pulmonary disease (COPD), chronic heart failure, AIDS, and chronic kidney disease**, among other conditions. Its prevalence is of around 5 to 15% in cardiac patients at end stage, rising up to 30%, in COPD and chronic kidney disease patients, and to **80% in patients with advanced cancer**. Cachexia symptoms include **pronounced weight loss**, due to both lean and fat mass wasting: **anorexia, malabsorption, nausea, asthenia, neuroendocrine changes, immune system function impairment, and disruption of energy metabolism**.

Despite its unquestionable relevance to the poorer outcome of treatment in disease and its high prevalence among patients, the syndrome is **still underdiagnosed and seldom treated**. Part of the difficulty in treating cachexia relies on the fact that, in the clinical setting, the syndrome is recognised solely in its **most advanced stages**, when therapy available to the present day is not able to fully reverse its symptoms.

# Motivation

Therefore, scientists and clinicians should focus on **identifying early changes**, as to intervene in a precocious manner. Taken together, the issue provides insights on the importance of **detecting early signs of inflammatory changes** in patients and examines the **mechanisms that act in concert**, inducing cachexia symptoms.

Hindawi Publishing Corporation  
Mediators of Inflammation  
Volume 2015, Article ID 536954, 2 pages  
<http://dx.doi.org/10.1155/2015/536954>



## *Editorial*

### **Inflammation in Cachexia**

**M. Seelaender,<sup>1</sup> A. Laviano,<sup>2</sup> S. Busquets,<sup>3</sup> G. P. Püschel,<sup>4</sup>  
T. Margaria,<sup>5</sup> and M. L. Batista Jr.<sup>6</sup>**

<sup>1</sup>*Cancer Metabolism Research Group, Institute of Biomedical Sciences, University of São Paulo, Avenida Professor Lineu Prestes 1524, 05508-900 São Paulo, SP, Brazil*

<sup>2</sup>*Department of Clinical Medicine, Sapienza University, Viale dell'Università 37, 00185 Rome, Italy*

<sup>3</sup>*Cancer Research Group, Department of Biochemistry and Molecular Biology, Biology Faculty, University of Barcelona, Avenida Diagonal 643, 08028 Barcelona, Spain*

<sup>4</sup>*Department of Nutritional Biochemistry, Institute of Nutritional Science, University of Potsdam, Arthur-Scheunert-Allee 114, 11614558 Nuthetal, Germany*









<sup>5</sup>*Lero, The Irish Software Research Center, University of Limerick, Limerick, Ireland*

<sup>6</sup>*Laboratory of Adipose Tissue Biology, Integrated Group of Biotechnology, University of Mogi das Cruzes, Avenida Dr. Cândido Xavier de Almeida Souza 200, Centro Cívico, 08780-911 Mogi das Cruzes, SP, Brazil*

# External Groups



**Table 1.** Remote associated groups and their geographical location and affiliation.

Group	Institute/Country
Alessandro Laviano, Maurizio Muscaritoli	Department of Clinical Medicine Sapienza University of Rome, Rome, Italy. 
Giorgio Trinchieri, Romina Goldzmid	Center for Cancer Research National Cancer Institute, Bethesda, Maryland USA 
Josep M. Argilés, Silvia Busquets	Cancer Research Group, Institut de Biomedicina, Univ. Barcelona Barcelona, Spain 
Nicolaas Deutz	Department of Health & Kinesiology Texas A&M University Bryan, Texas, USA 
Stephen Farmer	Department of Biochemistry Boston University School of Medicine Boston, MA, USA 
Gerhard Paul Piischel	Institute of Nutritional Science University of Potsdam Potsdam, Germany 
<i>Tiziana Margaria</i>	<i>Institute of Informatics University of Potsdam Potsdam, Germany</i> 
<i>Barry D. Floyd</i>	<i>California State Polytechnic University San Luis Obispo, CA, USA</i> 



Application Experts



Technical Experts

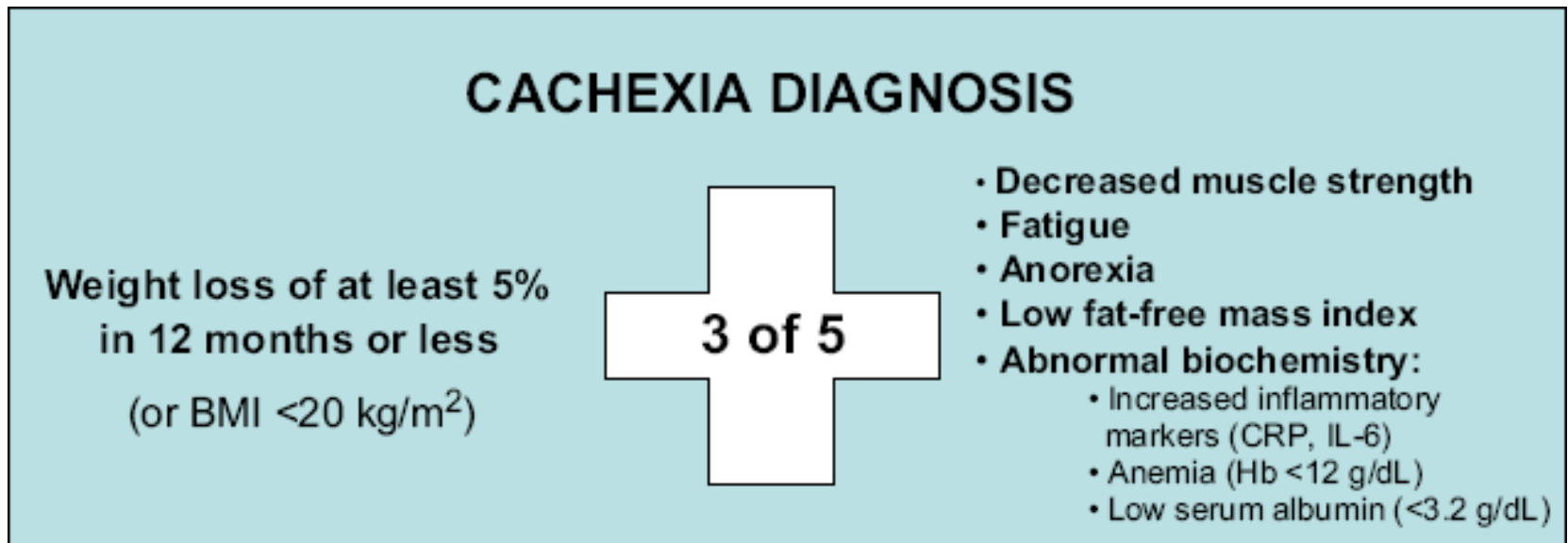
# Main Challenge

- Life sciences researchers are seldom IT professionals
  - they need to **work together efficiently**
  - with the data definition and management techniques that **complex and evolving experimental settings** require

# Definition of Cachexia

## Cachexia: A new definition

William J. Evans<sup>\*</sup>, John E. Morley<sup>a</sup>, Josep Argilés<sup>a</sup>,  
Connie Bales<sup>a</sup>, Vickie Baracos<sup>a</sup>, Denis Guttridge<sup>a</sup>,  
Aminah Jatoi<sup>a</sup>, Kamyar Kalantar-Zadeh<sup>a</sup>, Herbert Lochs<sup>a</sup>,  
Giovanni Mantovani<sup>a</sup>, Daniel Marks<sup>a</sup>, William E. Mitch<sup>a</sup>,  
Maurizio Muscaritoli<sup>a</sup>, Armine Najand<sup>a</sup>, Piotr Ponikowski<sup>a</sup>,  
Filippo Rossi Fanelli<sup>a</sup>, Morrie Schambelan<sup>a</sup>, Annemie Schols<sup>a</sup>,  
Michael Schuster<sup>a</sup>, David Thomas<sup>a</sup>, Robert Wolfe<sup>a</sup>, Stefan D. Anker<sup>a</sup>



# What **data** do we have?

- Questionnaire scores
- Anthropometric data
- Body composition images
- Brain images
- Histological images
- Immunodetection images
- Biochemical parameters in plasma
- Endocrine parameters in plasma
- Inflammatory parameters in plasma
- Biochemical parameters in tissues/organs
- Inflammatory parameters in tissues/organs
- Cell sorting and phenotyping spectra
- Chromatographic spectra
- Molecular parameters in cells
- Microarray analysis
- Gene sequencing
- Physiological parameters associated with the effect of exercise



# How to work with all these information?

- PDF Files
- Excel spreadsheets
- Everything is done manually



## Interdisciplinary:

Bio-chem lab  
Nutrition science  
Immunology  
Sports/rehab  
Oncology

Surgery  
Gastroenterology  
Psychology  
(Computer Science)  
(ED)

## Managing Processes

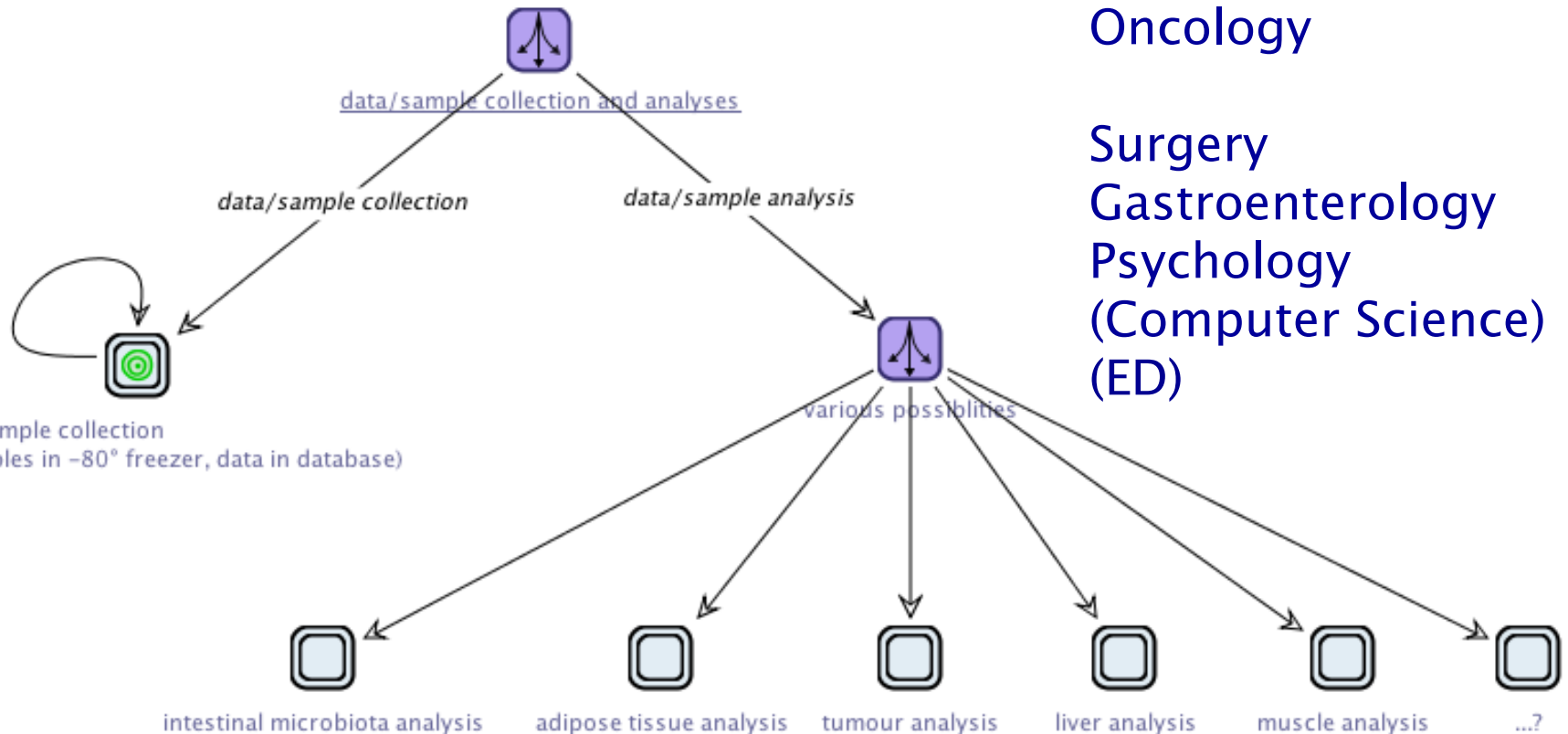


# Managing Processes

## Cachexia – the Big Picture (DAAD, FAPESP, CAPES)

**Interdisciplinary:**  
Bio-chem lab  
Nutrition science  
Immunology  
Sports/rehab  
Oncology


Surgery  
Gastroenterology  
Psychology  
(Computer Science)  
(ED)



# Stakeholder Requirements:

- **semiotically intuitive, graphical approach**
- **a framework that:**
  - is able to **manage complexity and change**
  - **helps to efficiently produce more reliable results**
  - appears **simple and intuitive** to its users

# The Patient Classification Spreadsheet

	A	B	C	D	E	F	G
1	<b>PATIENT'S INFORMATION</b>				<b>FIRST CRITERION - WEIGHT LOSS</b>		
2	Identification	Gender	Age (Years)		Weight variation	BMI (kg/m <sup>2</sup> )	Result
3	165A	Male	51		-10%	28,88	IN
4	Height (m)	Prev. weight (kg)	Current weight (kg)		Treatment	Hernia	
5	1,59	81	73				
6							
7	<b>SECOND CRITERION - WEIGHT STRENGTH</b>				<b>THIRD CRITERION - FATIGUE</b>		
8	Method	Score	Result		Method	Score	Result
9	Questionnaire (QLC-C30)	53,33333333	OUT		Questionnaire (QLC-C30)	33,33	IN
10	Answer 1	1			Answer 10	4	
11	Answer 2	4			Answer 12	3	
12	Answer 3	2			Answer 18	2	
13	Answer 4	3					
14	Answer 5	2					
15							
16	<b>FOURTH CRITERION - ANOREXIA</b>				<b>FIFTH CRITERION - FAT FREE MASS INDEX</b>		
17	Method	Score	Result		Method	Score	Result
18	Questionnaire (QLC-C30)	100,00	OUT		DEXA Scan	6,09	IN
19	Answer 13	1			Lean mass (kg)	15,4	
20							
21							
22	<b>SIXTH CRITERION - BIOCHEMICAL PARAMETERS</b>				<b>GROUP CLASSIFICATION</b>		<b>BARCODE</b>
23	Parameters	Concentration	Result		<b>CACHEXIA WITHOUT CANCER</b>		165ACWCENONE
24	C-Reactive protein (mg/l)	6,10	IN		<b>LEVEL OF EXCLUSION CRITERIA</b>		
25	IL-6 (pg/ml)	5,34					
26	Anemia - Hb (g/dl)	12,30					
27	Albumin (g/dl)	4,89					
28	Adapted from Evans, 2008				 0	<b>NONE</b>	

# Example: Excel Formula (Cell B9)

```
=WENN (UND (A9="Questionnaire (QLC-C30)";B10<>"";B11<>"";  
B12<>"";B13<>"";B14<>""); (1- (MITTELWERT (B10:B14) -1) /3) *100;  
WENN (UND (A9="HandgripTest";B3="Male";B10<44;B10<>""); "POOR";  
WENN (UND (A9="HandgripTest";B3="Male";B10>44; B10<>""); "GOOD";  
WENN (UND (A9="HandgripTest"; B3="Female";B10<23;B10<>""); "POOR";  
WENN (UND (A9="HandgripTest"; B3="Female";B10>22;B10<>"");  
"GOOD"; "---" ) ) ) )
```

- for regular use in large-scale research projects it has significant **drawbacks**:
  - for every patient the **data** has to be **entered manually** in a single spreadsheet  
→ also the result has to be **transferred manually**
  - Excel **formulas are complex** and not easily understandable
  - maintenance, modifications, and extensions are difficult and error-prone
  - **no statistics** on the data possible
  - **no unified database**

# The Patient Classification Spreadsheet

- complex **patient classification determines** which individuals belong to the different patient and control **groups**
- it is a **multifaceted evaluation** of a number of interdisciplinary **criteria**

We

- organized and simplified the patient classification process,
- made it **easily accessible** and **shareable** worldwide via a web application
- the integrated framework collects the data of patients in a **central repository** (easy to access and back up)

# Approach

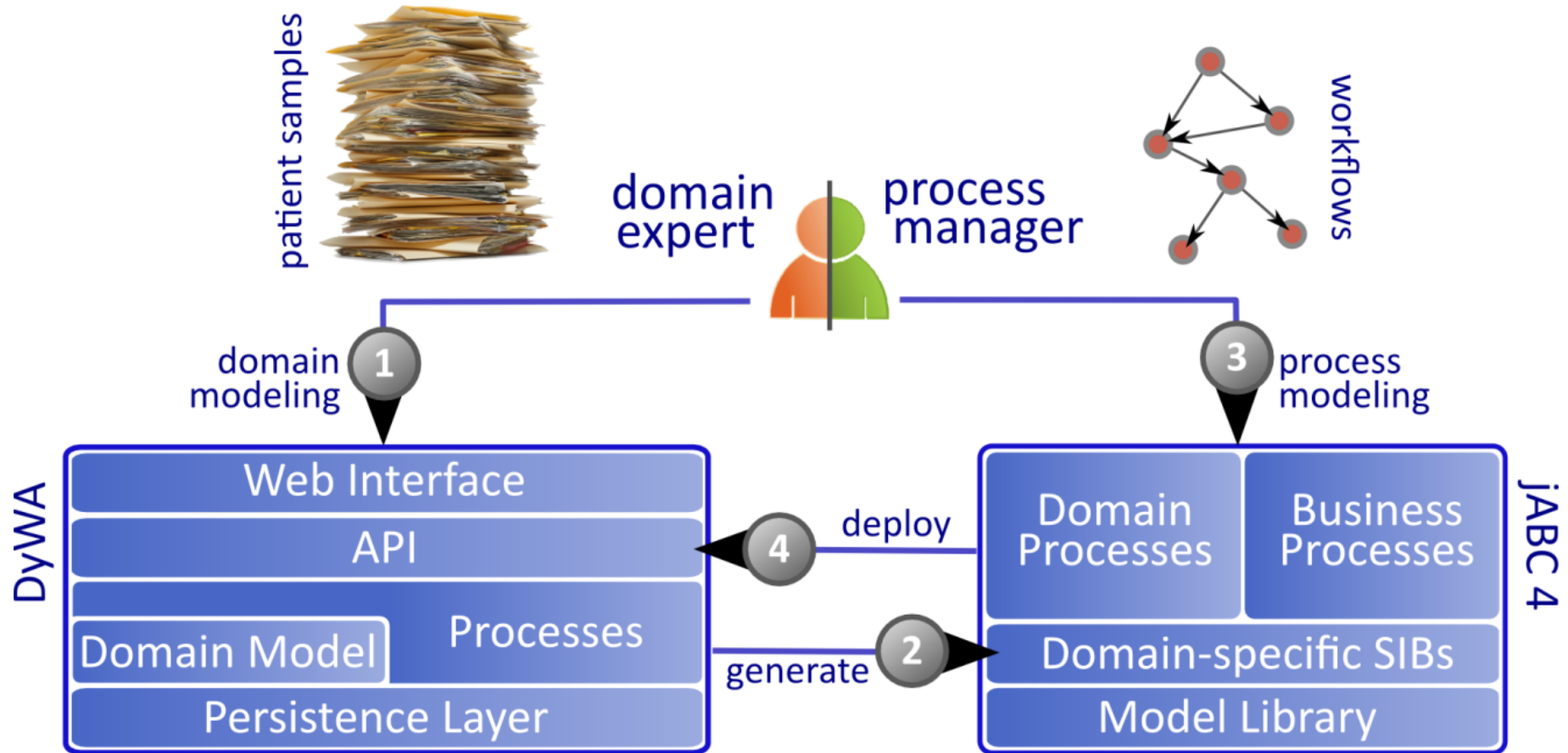
IDE for **co-design** and **co-evolution** of data and process models

- Data: Dynamic Web Application (**DyWA**),
- Behaviour: Java Application Building Center 4 (**jABC4**) modeling framework

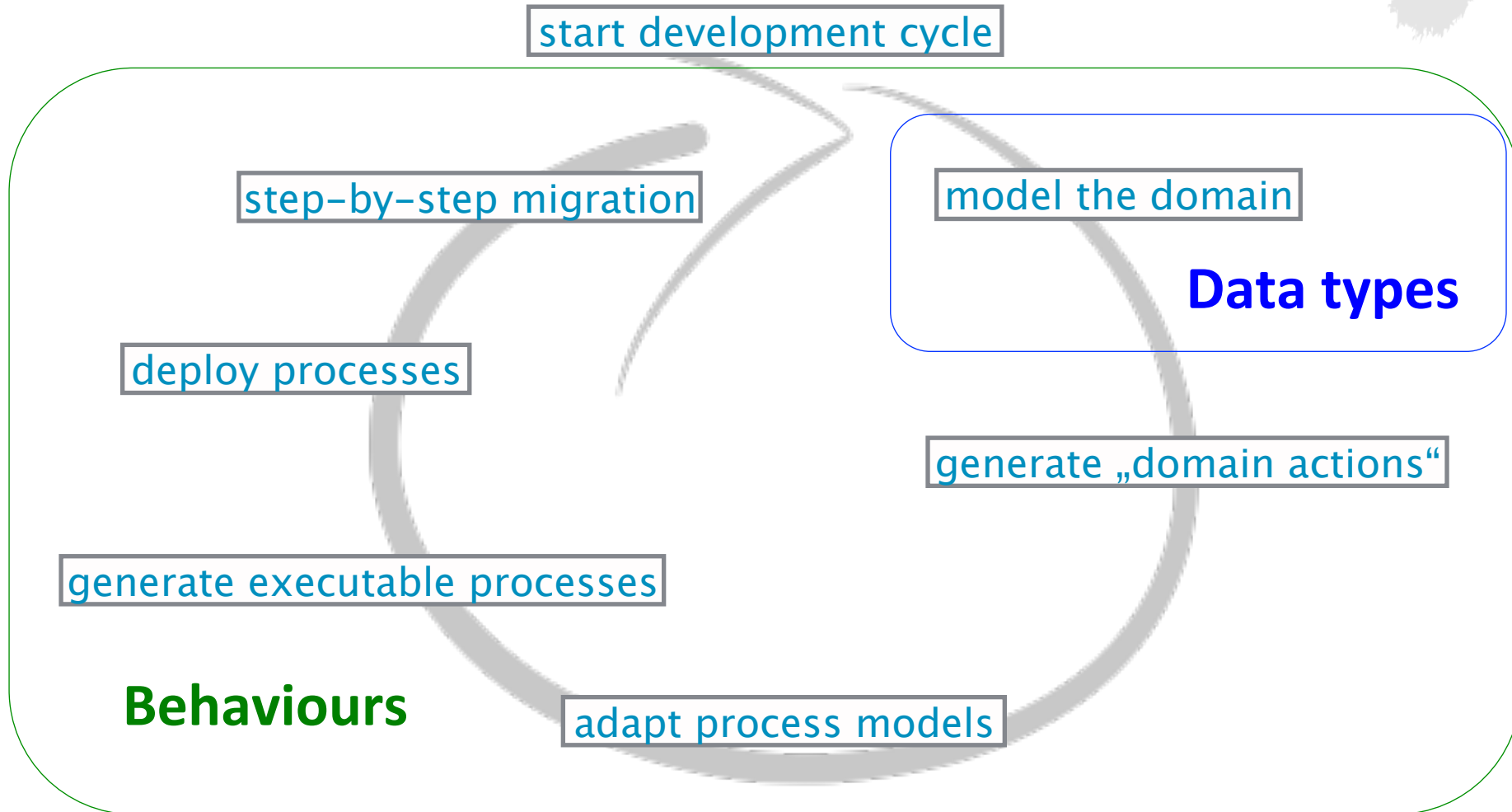
## Result:

easily executable domain-specific processes

# Co-Development with DyWA and jABC4



# Prototype-Driven Development (DevOps style)



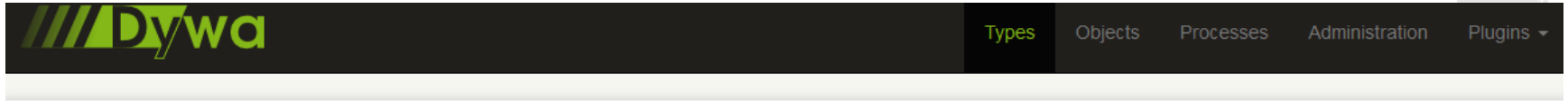


# DyWA: Data Type Definition

- DSL design approach = start by fixing the vocabulary of the domain
- definition of a set of domain-relevant “*things*” with their respective types
- DyWA provides **domain-independent** Java types (e.g. String, Integer) as initial type collection
- any **self-modeled** domain specific **type** (e.g. *Weight*) becomes directly available and can be used as a field or attribute of complex data types
- **CRUD** operations = **C**reate, **R**ead, **U**ppdate, **D**elate are automatically generated for every defined **type** and **field** by the DyWA

# DyWA: Data Type Definition

(Margaret Hamilton: who you **are**)



## Types

Available Types Add

Filter

- Answer\_enum\_values
- Gender\_enum\_values
- Height\_enum\_values
- Height\_type
- Identification
- Length\_enum\_values
- Length\_type
- Patient\_Information**
- Patient\_Questionary

### Patient\_Information

**Description**

**Inherits from**

**Parent of**

**Abstract** false

**Enumerable** false

**In Use** true

**Marked for deletion** false

**Used as** Patient\_Information

**Used in** There are no objects to display

**Fields**

Name	Type	In Use	As
identification	Identification	true	identification
gender	Gender_enum_values	true	gender
age	Integer number	true	age
height	Height_type	true	height
weight_previous	Weight_type	true	weight_previous
⋮	⋮	⋮	⋮

start development cycle

model the domain

# DyWA: Data Type Definition

**Edit type: Patient\_Information**

**Meta**

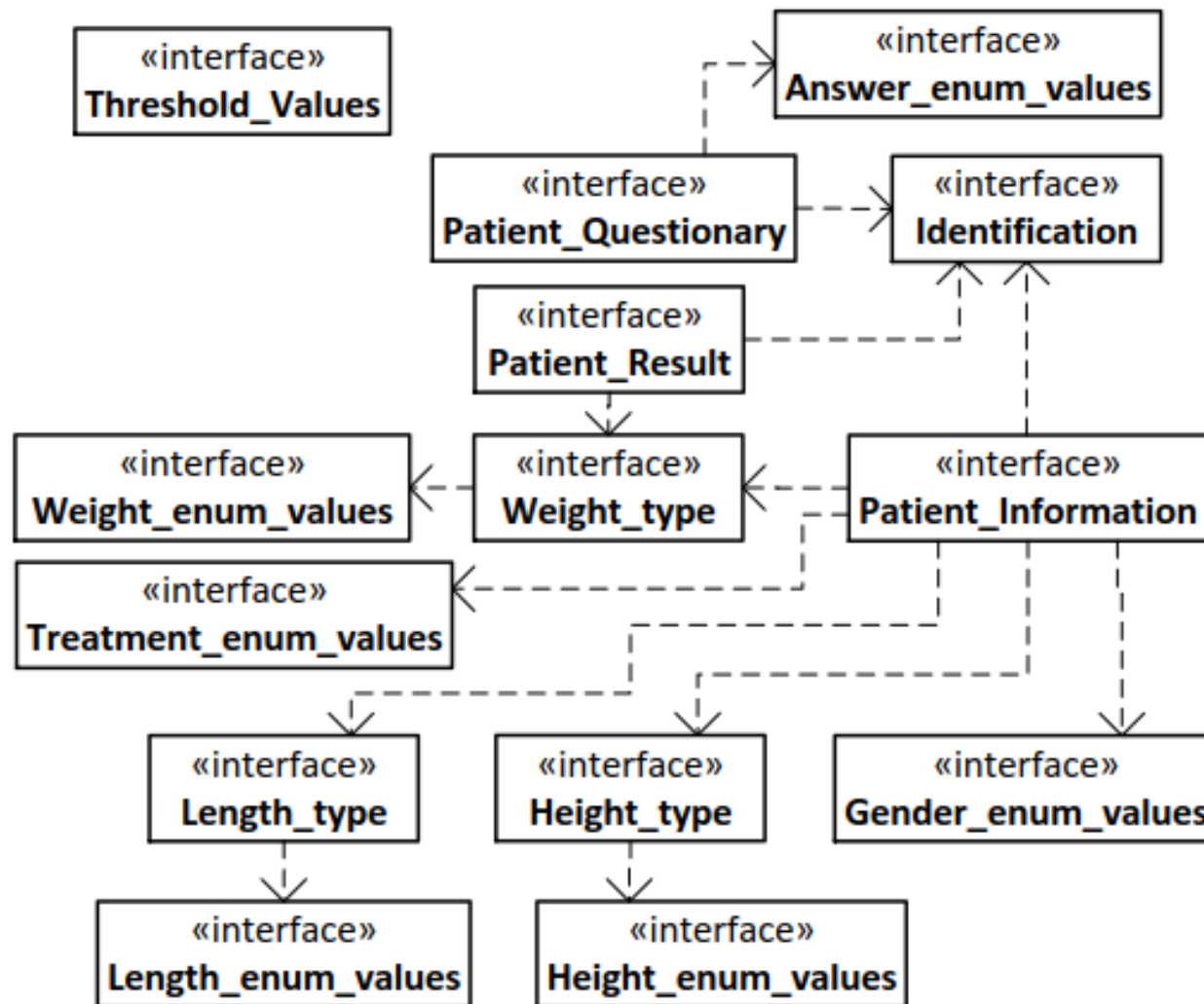
Name: Patient\_Information

**Fields of type**

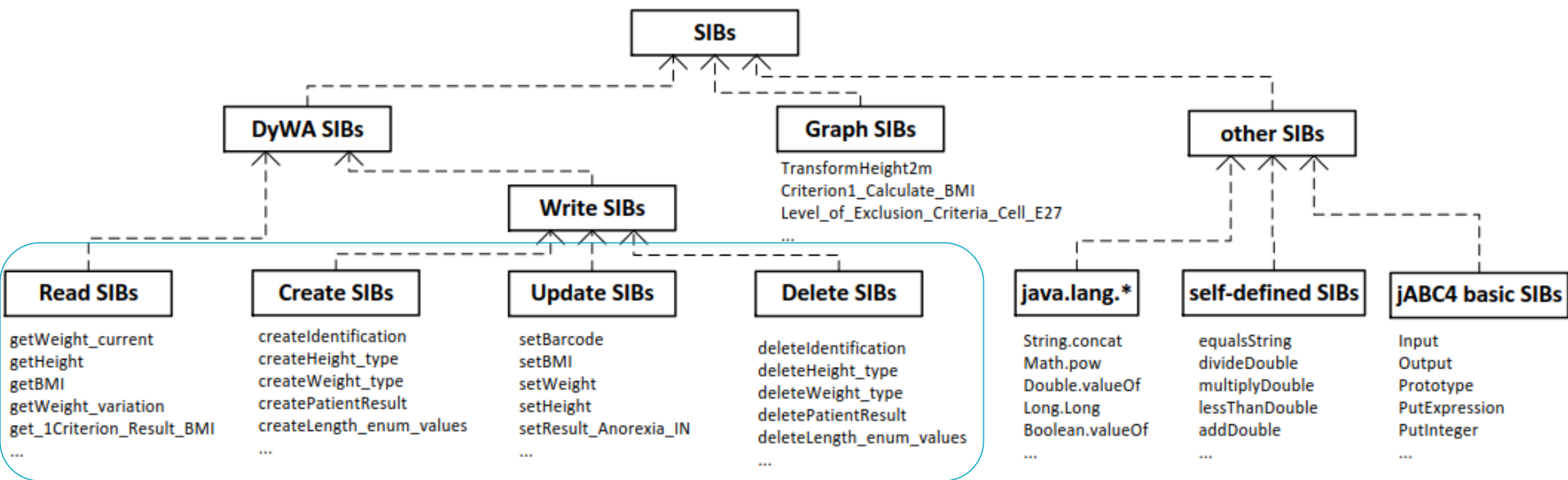
Index	Data Type	Value	Controls
1.	Identification	identification	↑ ↓ ↻ ✕ ▼
2.	Gender_enum_values	gender	↑ ↓ ↻ ✕ ▼
3.	Integer number	age	↑ ↓ ↻ ✕ ▼
4.	Height_type	height	↑ ↓ ↻ ✕ ▼
5.	Weight_type	weight_previous	↑ ↓ ↻ ✕ ▼
6.	Weight_type	weight_current	↑ ↓ ↻ ✕ ▼
⋮	⋮	⋮	⋮
15.	Floating-point number	albumin_in_g_per_dl	↑ ↓ ↻ ✕ ▼

Text [ ] +

# UML Class Diagram of the Domain Model

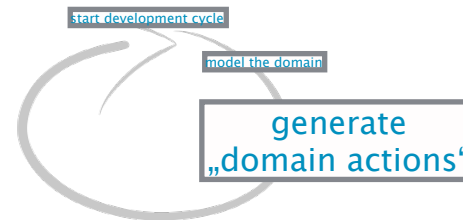


# Taxonomy of SIBs: the „Microservices“ (Margaret Hamilton: what you do)



CRUD SIBs automatically generated

“emerging” DSL for healthcare



# Process Modeling with jABC4

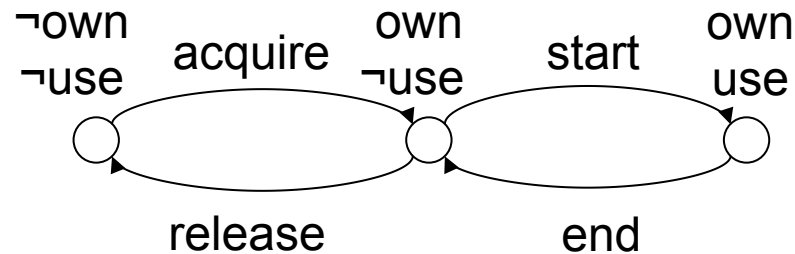
The screenshot displays the jABC4 software interface, divided into several panels:

- Top Panel:** Contains the menu bar (File, Edit, Project, SIB, Edge, Graph, View, Mode, Extras, Plugins, Help) and a toolbar with various icons. A circled '1' is placed over the toolbar.
- Left Panel (Projects):** Shows a tree view of the project structure. Under 'SIBs', there is a folder 'dywa' containing an 'entity' folder, which in turn contains 'Answer\_enum\_values'. This folder has sub-folders for 'read' and 'write', each containing several actions like 'listAnswer\_enum\_values', 'readAnswer\_enum\_values', 'readvalue', 'createAnswer\_enum\_values', 'deleteAnswer\_enum\_values', and 'writevalue'. A circled '2' is placed over this tree view.
- Bottom-Left Panel (SLG):** Shows configuration for a service. The 'Package' is 'de.ls5.dywa.generated.process' and the 'Name' is 'Criterion1\_Calculate\_BMI'. The 'Type' is set to 'service'. A circled '4' is placed over the 'Name' field. Below, a list of variables is shown: BMI (Double), Height\_type (Height\_type), Patient\_Information (Patient\_Informa), Patient\_Result (Patient\_Result), and Weight\_current (Weight\_type). The 'Interfaces' section is empty.
- Right Panel (Graph):** Displays a BPMN diagram for the 'Criterion1\_Calculate\_BMI' process. The process starts with a 'start' event (blue arrow) leading to a 'read weight\_current' activity (blue eye icon). This is followed by a 'TransformWeight2kg' activity (green G icon). Then, a 'read height\_type' activity (blue eye icon) leads to a 'TransformHeight2m' activity (green G icon). This is followed by a 'power(height\_in\_m, 2.0) = height\_to\_square' activity (green S icon) and a 'weight\_in\_kg / height\_to\_square = BMI' activity (green S icon). The process concludes with a 'write BMI' activity (orange eye icon) leading to a 'success' event (orange arrow). A circled '3' is placed over the 'read height\_type' activity. A large grey arrow on the right points from the diagram to a box labeled 'adapt process models'.

Additional annotations on the right side include:

- A box labeled 'start development cycle' with an arrow pointing to the start event.
- A box labeled 'model the domain' with an arrow pointing to the 'read height\_type' activity.
- A box labeled 'generate „domain actions“' with an arrow pointing to the 'TransformHeight2m' activity.

# Process Models: Kripke Transition System (KTS)



Let  $AP$  be a set of atomic propositions.

A **Kripke Transition System** over  $AP$  is a 4-tuple

$K=(S, Act, Trans, I)$  with:

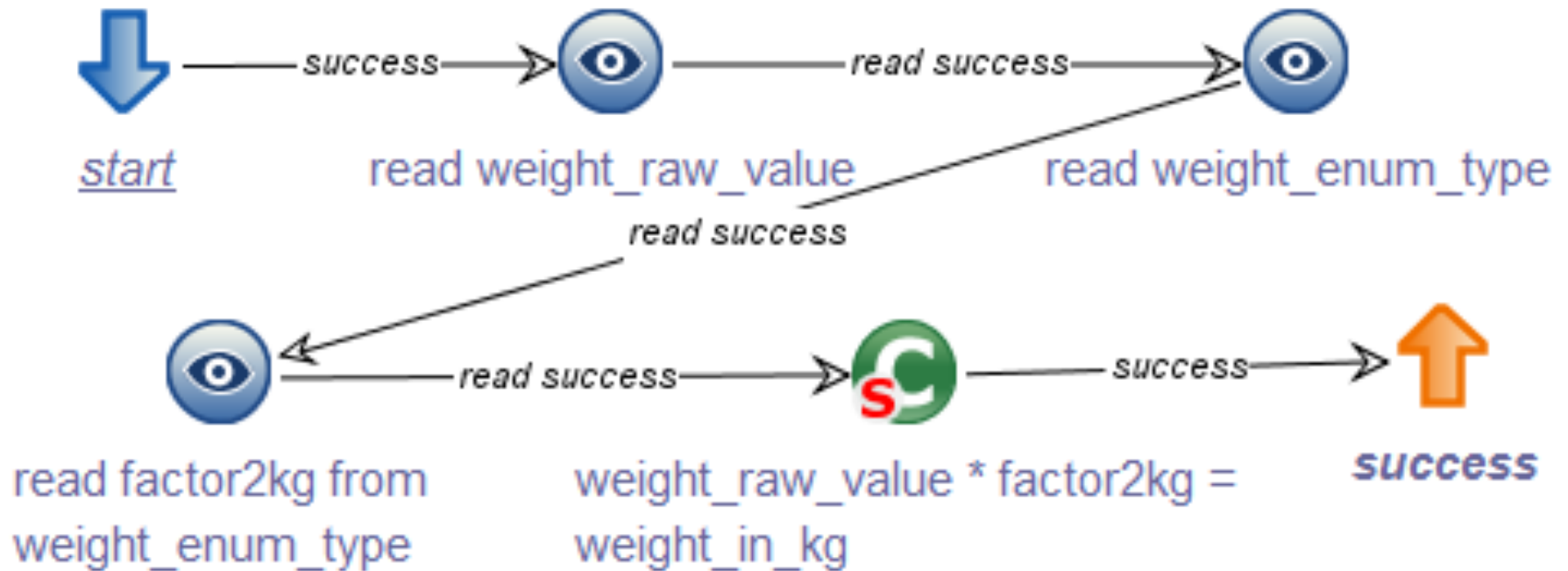
$S$  a set of states

$Act$  a set of actions

$Trans \subseteq S \times Act \times S$  a transition relation

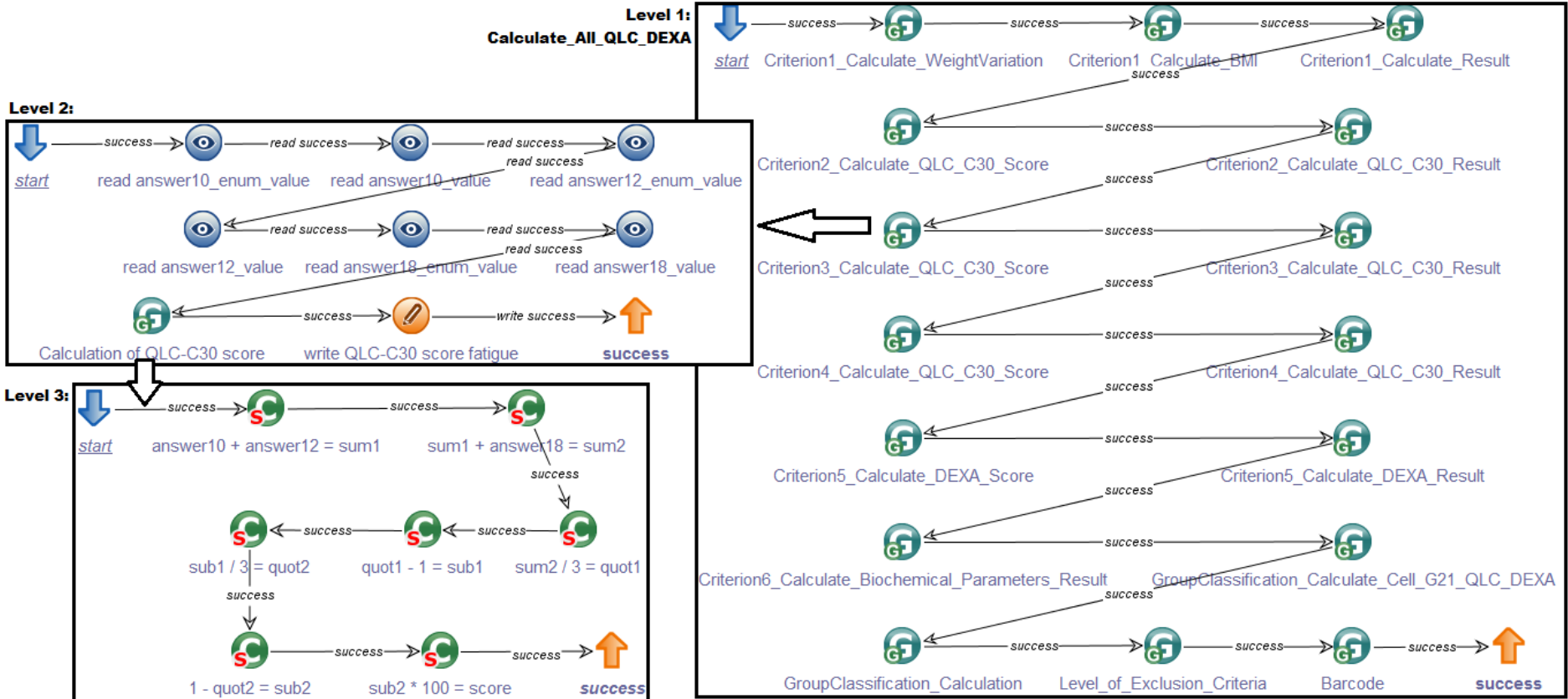
$I: S \rightarrow 2^{AP}$  an interpretation function

# Process Model: TransformWeight2kg





# Path through Graph Level Hierarchy

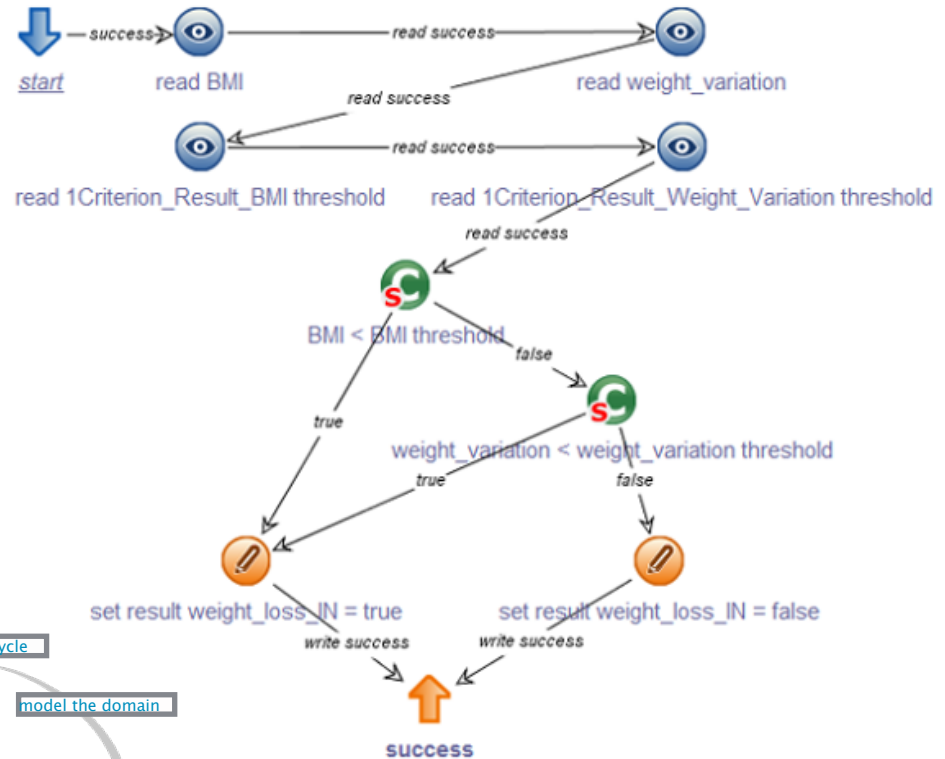


# Executable Processes deployed to DyWA

Available Processes
Filter
Barcode
Calculate_All_Handgrip_DEXA
Calculate_All_Handgrip_MUAMA
Calculate_All_QLC_DEXA
Calculate_All_QLC_MUAMA
Criterion1_Calculate_BMI
<b>Criterion1_Calculate_Result</b>
Criterion1_Calculate_WeightVariation
Criterion2_Calculate_HandgripTest_Result
Criterion2_Calculate_HandgripTest_Score
Criterion2_Calculate_QLC_C30_Result

Criterion1\_Calculate\_Result Execute

**Description**  
**Process Canvas**



deploy processes

generate executable processes

start development cycle

model the domain

generate „domain actions“

adapt process models

**Week 0  
(Pre-Training)**



adaptation

→ assess body composition



→ submaximal effort test  
to determine individual intensity  
(train always in this cardiac frequency range)



→ blood sampling



training always within the determined intensity (cardiac frequency)

**Week 1**



Monday-Friday:  
5 bouts of 3 min, 1 min between bouts  
record speed and cardiac frequency

**Week 2**



Monday-Friday:  
4 bouts of 5 min, 1 min between bouts  
record speed and cardiac frequency

**Week 3**



Monday-Friday:  
3 bouts of 8 min, 1 min between bouts  
record speed and cardiac frequency

→ submaximal test



→ blood sampling



**Week 4**



Monday-Friday:  
3 bouts of 10 min, 1 min between bouts  
record speed and cardiac frequency

**Week 5**



Monday-Friday:  
2 bouts of 15 min, 1 min between bouts  
record speed and cardiac frequency

**Week 6**



Monday-Friday:  
1 bout of 30 min  
record speed and cardiac frequency

→ submaximal test



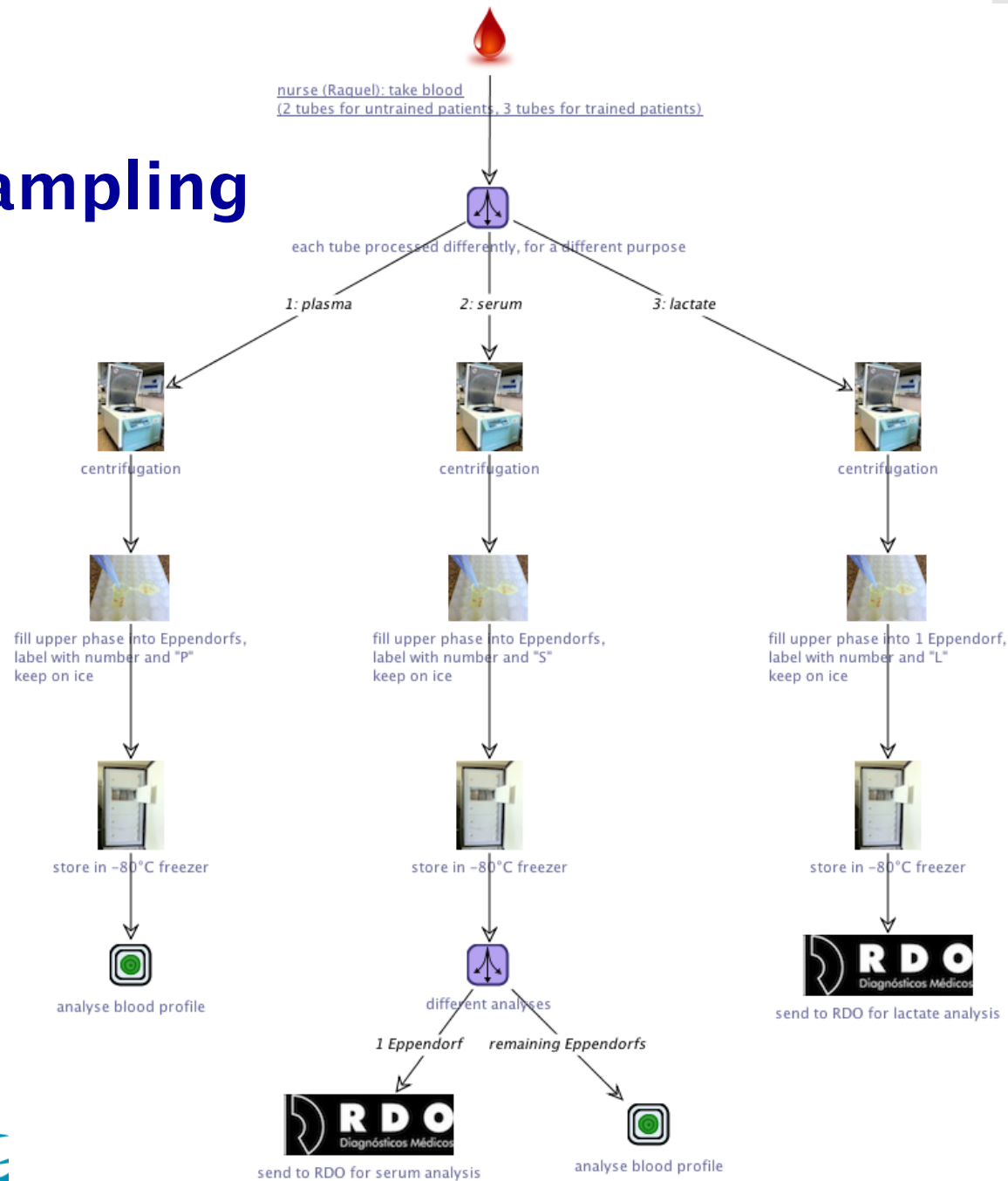
→ blood sampling











what to do with the data/protocols?

# Patient Training (treadmill, 6 weeks)

# Blood sampling



# Outcome: Incremental Modeling of Data and Processes

Group	Institute/Country
Alessandro Laviano, Maurizio Muscaritoli	Department of Clinical Medicine Sapienza University of Rome, Rome, Italy. 
Giorgio Trinchieri, Romina Goldzmid	Center for Cancer Research National Cancer Institute, Bethesda, Maryland USA 
Josep M. Argilés, Silvia Busquets	Cancer Research Group, Institut de Biomedicina, Univ. Barcelona Barcelona, Spain 
Nicolaas Deutz	Department of Health & Kinesiology Texas A&M University Bryan, Texas, USA 
Stephen Farmer	Department of Biochemistry Boston University School of Medicine Boston, MA, USA 
Gerhard Paul Püschel	Institute of Nutritional Science University of Potsdam Potsdam, Germany 
<i>Tiziana Margaria</i>	<i>Institute of Informatics University of Potsdam Potsdam, Germany</i> 
<i>Barry D. Floyd</i>	<i>California State Polytechnic University San Luis Obispo, CA, USA</i> 



Application Experts



Technical Experts

# Outcome:

## Incremental Modeling of Data and Processes

- patient classification process is now fully automated as a workflow
  - modeled with jABC4, and
  - integrated into a database provided by the DyWA
- **jABC4** = model-driven environment for designing the processes
- **DyWA** = meta-schema based data definition and management tool with standard relational database
- their interplay provides an integrated environment for **data and process modeling** along the XMDD paradigm
- supports a Service-oriented Continuous Engineering approach to the formalization and definition of a **domain-specific language** and **process landscape**

# eXtreme Model Driven Design



<http://cinco.scce.info/>



+



<http://hope.scce.info>



<http://dime.scce.info/>

# Limerick's vision







Thank you!

Questions?

Contact: [tiziana.margaria@lero.ie](mailto:tiziana.margaria@lero.ie)

# Conclusion

- integrated data and process modeling environment:
  - data collection,
  - data transformation,
  - automation,
  - reproducibility of results
- many **processes** have the potential to be **reused** by other health care applications, or even in other domains

# Conclusion

- **adaptations** to changing experimental setups are still possible
- **processes** are **immediately executable** and remain customizable
- this environment provides a significant step towards the large-scale applicability of a **formal** model-based and methods-supported, model-driven, generative IDE for scientists
- the IDE ensures that the modeling of domain-specific data types **and** processes using these components happens in one coherent system at a user-accessible level
- **result:** immediate availability, consistency and reproducibility of the outcomes, and the coherence and evolvability of the entire collection of data schema and processes

# Ongoing Work

- **access to data and processes** in the web application should be based on a proper **roles and rights management** (e.g. DIME)
- **provenance tracking and auditing** of all the data collected and accessed
  - to know **who did what when with which permissions**, and
  - to maintain truly **complete records** of experimental results